Pak. J. Statist.
2010, Vol. 26(1), 151-170

INCREASING POWER OF ROBUST TEST THROUGH PRE-TESTING IN MULTIPLE REGRESSION MODEL

Rossita M Yunus[1] and **Shahjahan Khan**

Department of Mathematics and Computing, Australian Centre for Sustainable Catchments,
University of Southern Queensland, Toowoomba, Q 4350, Australia.
Emails: yunus@usq.edu.au, khans@usq.edu.au

ABSTRACT

In this paper, we define (i) unrestricted (UT), (ii) restricted (RT) and (iii) pre-test (PTT) tests for testing the significance of a subset of the parameters of a multiple regression model when the remaining parameters are (i) completely unspecified, (ii) specified at fixed values or (iii) suspected at fixed values. The M-estimation methodology is used to formulate the three tests. The asymptotic distribution of the test statistics are used to derive the asymptotic power function of the tests. Analytical and graphical comparisons of the three tests are obtained by studying the power functions with respect to the size and power of the tests. The PTT shows a reasonable dominance over the other two tests asymptotically.

KEYWORDS

pre-test, asymptotic distribution, asymptotic power, M-estimation, contiguity, bivariate non-central chi-square.

2000 Mathematics Subject Classification: 62G20, 62G35, 62G10

1 Introduction

The multiple regression model is arguably the most commonly used statistical model in many real life problems. In many cases a large number of explanatory variables are included in the multiple regression model. Not all the explanatory variables contribute sig-

[1]on leave from Institute of Mathematical Sciences, Faculty of Sciences, University of Malaya, Malaysia.

nificantly to the prediction of the response variable. There are several methods available to exclude the non-significant explanatory variables from the model. In the context of testing hypotheses on any arbitrary subset of regression parameters, one may use the non-sample prior information on the explanatory variables to improve the power of the test on the coefficients of the remaining explanatory variables. Depending on the type of prior information various tests may be derived. It is important to investigate and compare the properties of the tests in order to select a test with the maximum power.

Let $X_i,\ i = 1,\dots,n$, be n observable response variables of a multiple regression model,

$$X_i = \beta' c_i + e_i, \tag{1.1}$$

where $\beta' = (\beta_1, \beta_2, \dots, \beta_p)$ is a p-dimensional row vector of unknown regression parameters, $c_i' = (c_{1i}, \dots,\ c_{pi})$ is a p-dimensional row vector of known real constants of the independent variables, e_i is the error term which is identically and independently distributed symmetric about 0 with a distribution function, $F_i,\ i = 1,\dots,n$. The vector of p-regression parameters can be expressed as $\beta' = (\beta_1',\ \beta_2')$ where β_1' is a sub-vector of order r and β_2' is a sub-vector of dimension s such that $r+s = p$. Similarly, partition c_i' as $(c_{i1}',\ c_{i2}')$ with $c_{i1}' = (c_{1i}, \dots,\ c_{ri})$ and $c_{i2}' = (c_{(r+1)i}, \dots, c_{pi})$.

Consider testing the significance of the sub-vector β_1 under three conditions on the values of the sub-vector β_2: (i) unspecified (ii) specified and fixed (iii) uncertain. For case (i), we want to test $H_0^\star : \beta_1 = 0$ against $H_A^\star : \beta_1 > 0$ with test function, ϕ_n^{UT}. This test is called the unrestricted test (UT). For case (ii), the test for testing of $H_0^\star : \beta_1 = 0$ against $H_A^\star : \beta_1 > 0$ with test function ϕ_n^{RT} is called the restricted test (RT). For case (iii), testing $H_0^{(1)} : \beta_2 = 0$ is recommended to remove the uncertainty of the suspicious values of $\beta_2 = 0$ before testing the significance of β_1. The testing on $H_0^{(1)} : \beta_2 = 0$ against $H_A^{(1)} : \beta_2 > 0$ with test function ϕ_n^{PT} is known as a pre-test (PT). If the null hypothesis of this pre-test is rejected, the UT is used to test $H_0^\star$, otherwise the RT is used. The ultimate test for testing $H_0^\star$ following a pre-testing on $H_0^{(1)}$ is defined as the pre-test test (PTT) and the test function is denoted by ϕ_n^{PTT}.

Many studies considered the estimation of parameters with non-sample prior information on the values of the parameters for various models. This includes commonly known preliminary test and Stein-type shrinkage estimators (see for instance, Khan and Saleh, 1997 and 2001, [1, 2]). However, in this paper we pursue the testing problem of the intercept parameter under non-sample prior information on the slope parameter. Some studies on the effect of the PT on the power of the PTT are found in literature for parametric cases [3, 4] as well as for the non-parametric cases [5, 6, 7]. Tamura (1965) [5] investigated the performance of the PTT for one sample and two samples non-parametric problems. Saleh and Sen (1982, 1983) [6, 7] proposed UT, RT and PTT based on rank test for simple linear model and simple multivariate model in two separate articles. Using sampling distribution theory of rank statistics, they developed the procedures to obtain the power functions for each test. Other than these two regression models, the UT, RT and PTT are also proposed for the parallelism model. As such, Lambert et al. (1985) [8] derived size and power of the UT, RT and PTT using the tests that are based on the least-squares (LS) estimators. Although the tests based on LS estimators do not depend on the assumptions of the underlying distribution of the error term, the LS estimates are identical to the maximum likelihood (ML) estimates when the distribution of the error term is assumed to be normally distributed. Both ML estimates and LS estimates are non robust with respect to deviation from the assumed (normal) distribution (c.f. [9, p.21]). So, it is suspected that the tests based on LS estimators are also non robust. The robust R-estimators are derived from the rank tests because rank tests are asymptotically distribution free under the null hypothesis (c.f. [10, p.281]). However, rank tests often preserve information about the order of the data but discard the actual values, thus overlook information that may have led to a better solution [11]. The most popular robust estimation method, however, is the M-estimation. The M-estimation method is applied to the actual data, hence, its statistical test does not suffer the same kind of lost of information as the rank test. One of the robust tests formulated using the M-estimation methodology is the M-test. The M-test is originally proposed by Sen (1982) [12] using the score function in the M-estimation methodology. It is expected that the M-test formulated by the M-estimation procedures inherits the robustness properties of the estimation method, so the test is less sensitive to departures from model assumptions. Sen (1982) [12] introduced the M-test for testing the significance of β_2 only. Recently, few unpublished papers [13] have used M-test for the UT, RT and PTT in the regression model to investigate the performance of the tests.

In this paper, the M-tests for the UT, RT and PTT are proposed for the multiple regression model. To author's knowledge, no research has been done in investigating the performance of the UT, RT and PTT for multiple regression model. There is no article found in literature proposing UT, RT and PTT for this model based on rank tests. So, this paper is the first attempt to the study of comparing the performance of UT, RT and PTT for the multiple regression model. The asymptotic distribution theory of the test statistics that are based on the score function in the M-estimation methodology developed by Jurečková (1977) [14] and Jurečková and Sen (1996) [9, Ch. 5] is used in this paper. Although the asymptotic results of [9, Ch.5] are used in deriving the distribution of the proposed test statistics; here these results are adopted for a different model in the context of testing after pre-test.

The investigations on the comparisons of the UT, RT and PTT for simple multivariate model [7] and parallelism model [8] are limited to analytical discussion only; the computational comparisons of the UT, RT and PTT are not provided in these papers. Perhaps, the computational comparison of the UT, RT and PTT could not be given due to the nonexistence tool to compute the power functions at that time. To compute the power of the PTT, the bivariate integral of the non-central chi-square distribution is required. However, the proposed bivariate non-central chi-square distribution in literature at the time their papers were published are very complicated and not practical for computation. In this paper, we refer to Yunus and Khan (2009b) [15] for the computation of the bivariate integral of the non-central chi-square distribution. For simple multivariate model and parallelism model, according to Saleh and Sen (1983) [7] and Lambert et al. (1985) [8], the power of the PTT may be between those of the UT and RT. However, this statement is not clearly supported by arguments in their papers probably due to the complicated form of the bivariate non-central chi-square distribution that they used in the papers. Although multiple regression model is considered in this paper, we obtain quite similar findings as the simple multivariate and parallelism models because the test statistic for the PTT is bivariate noncentral chi-square distribution for all of these models. We discussed their statement in the findings of this paper and support our findings with clear arguments and through simulated studies.

Along with some preliminary notions, the method of M-estimation is presented in Section 2. The UT, RT and PTT are defined in Section 3. In Section 4, the asymptotic distribution of the proposed test statistics are derived. These distributions are used to obtain the power functions of the tests in Section 5. The analytical comparisons of the UT, RT and PTT are also given in Section 5. The comparisons of the power function of the UT, RT and PTT through simulation example are provided in Section 6. The final section presents discussions and concluding remarks.

2 The M-estimation

Given an absolutely continuous function $\rho : \Re \to \Re$, M-estimator of β is defined as the solution of minimizing the objective function

$$\sum_{i=1}^{n} \rho\left(\frac{X_i - \beta' c_i}{S_n}\right) \tag{2.1}$$

with respect to $\beta \in \Re_p$. Here S_n is an appropriate scale statistic for some functional $S = S(F) > 0$. If F is $N(0,\sigma^2)$, $S_n = MAD/0.6745$ is an estimate of $S = \sigma$, where MAD is the mean absolute deviation ([16, p.78], [17, p.387]). If $\psi = \rho'$, then the M-estimator of β is the solutions of the system of equations,

$$\sum_{i=1}^{n} c_i \psi\left(\frac{X_i - \beta' c_i}{S_n}\right) = 0. \tag{2.2}$$

For any r and s dimensional column vectors, t_1 and t_2 ($r,s \in \Re$), consider the statistics below

$$M_{n_1}(t_1,t_2) = \sum_{i=1}^{n} c_{i1} \psi\left(\frac{X_i - t_1' c_{i1} - t_2' c_{i2}}{S_n}\right), \tag{2.3}$$

$$M_{n_2}(t_1,t_2) = \sum_{i=1}^{n} c_{i2} \psi\left(\frac{X_i - t_1' c_{i1} - t_2' c_{i2}}{S_n}\right). \tag{2.4}$$

For a nondecreasing $\psi : \Re \to \Re$, let $\tilde{\beta}_2$ be the constrained M-estimator of β_2 when $\beta_1 = 0$, that is, $\tilde{\beta}_2$ is the solution of $M_{n_2}(0,t_2) = 0$ and it may be conveniently be expressed as

$$\tilde{\beta}_2 = [\sup\{t_2 : M_{n_2}(0,t_2) > 0\} + \inf\{t_2 : M_{n_2}(0,t_2) < 0\}]/2 \tag{2.5}$$

(cf. [12]). Note that for nondecreasing ψ function, $M_{n_2}(0,t_2)$ is decreasing as t_2 is increasing (c.f. [9, p.85]). Similarly, let $\tilde{\beta}_1$ be the constrained M-estimator of β_1 when $\beta_2 = 0$, that is, $\tilde{\beta}_1$ is the solution of $M_{n_1}(t_1,0) = 0$ and conveniently be expressed as

$$\tilde{\beta}_1 = [\sup\{t_1 : M_{n_1}(t_1,0) > 0\} + \inf\{t_1 : M_{n_1}(t_1,0) < 0\}]/2. \tag{2.6}$$

Theorem 1. *Given the asymptotic properties of $M_{n_1}(\cdot,\cdot)$ and $M_{n_2}(\cdot,\cdot)$ in equations (A.1), (A.2) and (A.3) in the Appendix A, asymptotically,*

(i) $\quad n^{-\frac{1}{2}} M_{n_1}(0,\tilde{\beta}_2) \xrightarrow{d} N_r(0,\sigma_0^2 Q_1^\star)$ *under* $H_0^\star : \beta_1 = 0,$ $\quad$ (2.7)

(ii) $\quad n^{-\frac{1}{2}} M_{n_2}(\tilde{\beta}_1,0) \xrightarrow{d} N_s(0,\sigma_0^2 Q_2^\star)$ *under* $H_0^{(1)} : \beta_2 = 0,$ $\quad$ (2.8)

where $Q_1^\star = Q_{11} - Q_{12} Q_{22}^{-1} Q_{21}$ *and* $Q_2^\star = Q_{22} - Q_{21} Q_{11}^{-1} Q_{12}$.

Here, $N_r(\cdot\,,\cdot\,)$ represents an r-variate normal distribution with appropriate parameters. Let $\sigma_0^2 = \int_{-\infty}^{\infty} \psi^2\left(\frac{X-\beta' c}{S}\right) dF(X - \beta' c)$. Here σ_0^2 is the second moment of $\psi(\cdot)$ while the first moment is zero by assuming F is symmetrically distributed at 0 and ψ is a skew symmetric function. Take $Q_{11} = \lim_{n\to\infty} \frac{1}{n} Q_{n_{11}} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} c_{i1} c_{i1}'$, $Q_{12} = \lim_{n\to\infty} \frac{1}{n} Q_{n_{12}} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} c_{i1} c_{i2}'$, $Q_{21} = \lim_{n\to\infty} \frac{1}{n} Q_{n_{21}} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} c_{i2} c_{i1}'$ and $Q_{22} = \lim_{n\to\infty} \frac{1}{n} Q_{n_{22}} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} c_{i2} c_{i2}'$. Assume that $|Q_{11}| \neq 0$, $|Q_{22}| \neq 0$, $|Q_1^\star| \neq 0$ and $|Q_2^\star| \neq 0$.

3 The UT, RT and PTT

3.1 The unrestricted test (UT)

If β_2 is unspecified, ϕ_n^{UT} is the test function of $H_0^\star : \beta_1 = 0$ against $H_A^\star : \beta_1 > 0$. Under $H_0^\star$, $X_i = \beta_2' c_{i2} + e_i$. We consider test statistic

$$T_n^{UT} = \frac{M_{n_1}(0,\tilde{\beta}_2)' Q_{n_1}^{\star\,-1} M_{n_1}(0,\tilde{\beta}_2)}{S_n^{(1)^2}},$$

where $\tilde{\beta}_2$ (given in equation (2.5)) is a constrained M-estimator of β_2 under $H_0^\star$. It follows from equation (2.8) that T_n^{UT} is χ_r^2 (chi-square distribution with r degrees of freedom) under $H_0^\star$ as $n \to \infty$, with $Q_{n_1}^\star = Q_{n_{11}} - Q_{n_{12}} Q_{n_{22}}^{-1} Q_{n_{21}}$ and $S_n^{(1)^2} = \sum \psi^2 \left(\frac{X_i - \tilde{\beta}_2' c_{i2}}{S_n} \right) / n$.

Let ℓ_{n,α_1}^{UT} be the critical value of T_n^{UT} at the α_1 level of significance. So, for the test function $\phi_n^{UT} = I(T_n^{UT} > \ell_{n,\alpha_1}^{UT})$, the power function of the UT becomes $\Pi_n^{UT}(\beta_1) = E(\phi_n^{UT} | \beta_1) = P(T_n^{UT} > \ell_{n,\alpha_1}^{UT} | \beta_1)$, where $I(A)$ is an indicator function of the set A. It takes value 1 if A occurs, otherwise it is 0.

3.2 The restricted test (RT)

If $\beta_2 = 0$, ϕ_n^{RT} is the test function for testing $H_0^\star : \beta_1 = 0$ against $H_A^\star : \beta_1 > 0$. The proposed test statistic is

$$T_n^{RT} = \frac{M_{n_1}(0,0)' Q_{n_{11}}^{\ -1} M_{n_1}(0,0)}{S_n^{(2)^2}}.$$

It follows from equation (A.3) that for large n, $T_n^{RT} \xrightarrow{d} \chi_r^2$ under $H_0 : \beta_1 = 0, \beta_2 = 0$ where $S_n^{(2)^2} = \sum \psi^2 \left(\frac{X_i}{S_n} \right) / n$. Again, let ℓ_{n,α_2}^{RT} be the critical value of T_n^{RT} at the α_2 level of significance. So, for the test function $\phi_n^{RT} = I(T_n^{RT} > \ell_{n,\alpha_2}^{RT})$, the power function of the RT becomes $\Pi_n^{RT}(\beta_1) = E(\phi_n^{RT} | \beta_1) = P(T_n^{RT} > \ell_{n,\alpha_2}^{RT} | \beta_1)$.

3.3 The pre-test (PT)

For the preliminary test on the slope, ϕ_n^{PT} is the test function for testing $H_0^{(1)} : \beta_2 = 0$ against $H_A^{(1)} : \beta_2 > 0$. Under $H_0^{(1)}$, $X_i = \beta_1' c_{i1} + e_i$. The proposed test statistic is

$$T_n^{PT} = \frac{M_{n_2}(\tilde{\beta}_1, 0)' Q_{n_2}^{\star\,-1} M_{n_2}(\tilde{\beta}_1, 0)}{S_n^{(3)^2}},$$

where $\tilde{\beta}_1$ (given in equation (2.6)) is a constrained M-estimator of β_1. It follows from equation (2.7) that $T_n^{PT} \xrightarrow{d} \chi_s^2$ under $H_0^{(1)}$, where $Q_{n_2}^{\star} = Q_{n_{22}} - Q_{n_{21}} Q_{n_{11}}^{-1} Q_{n_{12}}$ and ${S_n^{(3)}}^2 = \sum \psi^2 \left(\frac{X_i - \tilde{\beta}_1' c_{i1}}{S_n} \right) / n$.

3.4 The pre-test test (PTT)

Let ϕ_n^{PTT} be the test function for testing $H_0^{(1)}$ following a pre-test on β. Since the PTT is a choice between RT and UT, define,

$$\phi_n^{PTT} = I[(T_n^{PT} < \ell_{n,\alpha_3}^{PT}, T_n^{RT} > \ell_{n,\alpha_2}^{RT}) \text{ or } (T_n^{PT} > \ell_{n,\alpha_3}^{PT}, T_n^{UT} > \ell_{n,\alpha_1}^{UT})], \tag{3.1}$$

where ℓ_{n,α_3}^{PT} is the critical value of T_n^{PT} at the α_3 level of significance. The power function of the PTT is given by

$$\Pi_n^{PTT}(\beta_1) = E(\phi_n^{PTT} | \beta_1) \tag{3.2}$$

and the size of the PTT is obtained by substituting $\beta_1 = 0$ in equation (3.2).

4 Asymptotic distribution of UT, RT, PT and PTT

In this section, the asymptotic distributions of UT, RT, PT and PTT are derived under local alternative hypotheses, $\{K_n\}$ (see below). These distributions are essential to obtain the power functions of the UT, RT and PTT. To derive the power function of the PTT, we require to find the joint distributions of $\left[T_n^{UT}, T_n^{PT}\right]$ and $\left[T_n^{RT}, T_n^{PT}\right]$.

Theorem 2. *Let $\{K_n\}$ be a sequence of local alternative hypotheses, where*

$$K_n : (\beta_1, \beta_2) = (n^{-\frac{1}{2}} \lambda_1, n^{-\frac{1}{2}} \lambda_2), \tag{4.1}$$

with $\lambda_1 = n^{\frac{1}{2}} \beta_1 > 0$ and $\lambda_2 = n^{\frac{1}{2}} \beta_2 > 0$ are (fixed) real numbers. Under $\{K_n\}$, asymptotically,

(i)

$$\begin{bmatrix} n^{-\frac{1}{2}} M_{n_1}(0, \tilde{\beta}_2) \\ n^{-\frac{1}{2}} M_{n_2}(\tilde{\beta}_1, 0) \end{bmatrix} \xrightarrow{d} N_p \left[\begin{pmatrix} \gamma Q_1^{\star} \lambda_1 \\ \gamma Q_2^{\star} \lambda_2 \end{pmatrix}, \sigma_0^2 \begin{pmatrix} Q_1^{\star} & Q_{12}^{\star} \\ Q_{21}^{\star} & Q_2^{\star} \end{pmatrix} \right], \tag{4.2}$$

(ii)

$$\begin{bmatrix} n^{-\frac{1}{2}} M_{n_1}(0, 0) \\ n^{-\frac{1}{2}} M_{n_2}(\tilde{\beta}_1, 0) \end{bmatrix} \xrightarrow{d} N_p \left[\begin{pmatrix} \gamma(Q_{11}\lambda_1 + Q_{12}\lambda_2) \\ \gamma Q_2^{\star} \lambda_2 \end{pmatrix}, \sigma_0^2 \begin{pmatrix} Q_{11} & 0 \\ 0 & Q_2^{\star} \end{pmatrix} \right], \tag{4.3}$$

where $Q_{12}^{\star} = Q_{12}Q_{22}^{-1}Q_{21}Q_{11}^{-1}Q_{12} - Q_{12}$, $Q_{21}^{\star} = Q_{21}Q_{11}^{-1}Q_{12}Q_{22}^{-1}Q_{21} - Q_{21}$ *and* $\gamma = \frac{1}{S} \int_{-\infty}^{\infty} \psi' \left(\frac{X-\beta' c}{S} \right) dF(X - \beta' c)$.

Theorem 3. *Under* $\{K_n\}$, *asymptotically* (T_n^{RT}, T_n^{PT}) *are independently distributed as bivariate noncentral chi-square distribution with* (r,s) *degrees of freedom and* (T_n^{UT}, T_n^{PT}) *are distributed as correlated bivariate noncentral chi-square distribution with* (r,s) *degrees of freedom and noncentrality parameters,*

$$\theta^{UT} = \frac{\gamma^2}{\sigma_0^2}(\lambda_1' Q_1^{\star} \lambda_1), \tag{4.4}$$

$$\theta^{RT} = \frac{\gamma^2}{\sigma_0^2}(\lambda_1' Q_{11} \lambda_1 + \lambda_1' Q_{12} \lambda_2 + \lambda_2' Q_{21} \lambda_1 + \lambda_2' Q_{21} Q_{11}^{-1} Q_{12} \lambda_2), \tag{4.5}$$

$$\theta^{PT} = \frac{\gamma^2}{\sigma_0^2}(\lambda_2' Q_2^{\star} \lambda_2). \tag{4.6}$$

Proof. The proof of this theorem is directly obtained using Theorem 2 and Theorem 1.4.1 of Muirhead (1982) [18]. □

5 Asymptotic properties for UT, RT and PTT

Using results in Section 4, under $\{K_n\}$, the asymptotic power function for the UT is

$$\Pi^{UT}(\lambda_1, \lambda_2) = \lim_{n\to\infty} \Pi_n^{UT}(\lambda_1, \lambda_2) = \lim_{n\to\infty} P(T_n^{UT} > \ell_{n,\alpha_1}^{UT} | K_n) = 1 - G_r(\chi^2_{r,\alpha_1}; \theta^{UT}), \tag{5.1}$$

the asymptotic power function for the RT is

$$\Pi^{RT}(\lambda_1, \lambda_2) = \lim_{n\to\infty} \Pi_n^{RT}(\lambda_1, \lambda_2) = \lim_{n\to\infty} P(T_n^{RT} > \ell_{n,\alpha_2}^{RT} | K_n) = 1 - G_r(\chi^2_{r,\alpha_2}; \theta^{RT}), \tag{5.2}$$

and the asymptotic power function for the PT is

$$\Pi^{PT}(\lambda_1, \lambda_2) = \lim_{n\to\infty} \Pi_n^{PT}(\lambda_1, \lambda_2) = \lim_{n\to\infty} P(T_n^{PT} > \ell_{n,\alpha_3}^{PT} | K_n) = 1 - G_s(\chi^2_{s,\alpha_3}; \theta^{PT}), \tag{5.3}$$

where $G_k(\chi^2_{k,\alpha_\nu}; \theta^h)$ is the cumulative density function of the noncentral chi-square distribution with k degrees of freedom (d.f) and noncentrality parameter θ^h in which h is any of the UT, RT and PTT. The level of significance, $\alpha_\nu, \nu = 1,2,3$ is chosen together with the critical values ℓ_{n,α_ν}^h for the UT, RT and PT. Here, $\chi^2_{k,\alpha}$ is the upper $100\alpha\%$ critical value of a central chi-square distribution and $\ell_{n,\alpha_1}^{UT} \to \chi^2_{r,\alpha_1}$ under $H_0^{\star}$, $\ell_{n,\alpha_2}^{RT} \to \chi^2_{r,\alpha_2}$ under H_0 and $\ell_{n,\alpha_3}^{PT} \to \chi^2_{s,\alpha_3}$under $H_0^{(1)}$.

When $\theta^{RT} \geq \theta^{UT}$, the asymptotic size of the RT is larger than that of the UT but the asymptotic power of the UT is smaller than that of the RT. For testing $H_0^\star$ following a pre-test on β_2, using equation (3.1) and the results in Section 4, the asymptotic power function for the PTT under $\{K_n\}$ is given by

$$\begin{aligned}
&\Pi^{PTT}(\lambda_1,\lambda_2)\\
=\;& \lim_{n\to\infty} P(T_n^{PT} \leq \ell_{n,\alpha_3}^{PT}, T_n^{RT} > \ell_{n,\alpha_2}^{RT} | K_n) + \lim_{n\to\infty} P(T_n^{PT} > \ell_{n,\alpha_3}^{PT}, T_n^{UT} > \ell_{n,\alpha_1}^{UT} | K_n)\\
=\;& G_s(\chi^2_{s,\alpha_3};\theta^{PT})\{1 - G_r(\chi^2_{r,\alpha_2};\theta^{RT})\} + \int_{\chi^2_{r,\alpha_1}}^{\infty}\int_{\chi^2_{s,\alpha_3}}^{\infty} \phi^\star(w_1,w_2)dw_1dw_2, \qquad (5.4)
\end{aligned}$$

where $\phi^\star(\cdot)$ is the density function of a bivariate noncentral chi-square distribution. It is observed that $G_s(\chi^2_{s,\alpha_3};\, \theta^{PT})$ is decreasing as the value of θ^{PT} is increasing and $1 - G_r(\chi^2_{r,\alpha_2};\theta^{RT})$ is increasing as the value of θ^{RT} is increasing.

The probability integral in (5.4) is given by

$$\begin{aligned}
&\int_{\chi^2_{r,\alpha_1}}^{\infty}\int_{\chi^2_{s,\alpha_3}}^{\infty} \phi^\star(w_1,w_2)dw_1dw_2\\
=\;& \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{\delta_1=0}^{\infty}\sum_{\delta_2=0}^{\infty} (1-\rho^2)^{(r+s)/2} \frac{\Gamma(\frac{r}{2}+j)}{\Gamma(\frac{r}{2})j!} \frac{\Gamma(\frac{s}{2}+k)}{\Gamma(\frac{s}{2})k!} \rho^{2(j+k)}\\
&\times \left[1-\gamma^\star\left(\frac{r}{2}+j+\delta_1, \frac{\chi^2_{r,\alpha_1}}{2(1-\rho^2)}\right)\right]\left[1-\gamma^\star\left(\frac{s}{2}+k+\delta_2, \frac{\chi^2_{s,\alpha_3}}{2(1-\rho^2)}\right)\right]\\
&\times \frac{e^{-\theta^{UT}/2}(\theta^{UT}/2)^{\delta_1}}{\delta_1!} \frac{e^{-\theta^{PT}/2}(\theta^{PT}/2)^{\delta_2}}{\delta_2!}, \qquad (5.5)
\end{aligned}$$

with (r,s) degrees of freedom, noncentrality parameters, θ^{UT} and θ^{PT} and correlation coefficient, $-1 < \rho < 1$. Here, $\gamma^\star(v,d) = \int_0^d x^{v-1}e^{-x}/\Gamma(v)dx$ is the incomplete gamma function. For details on the evaluation of the bivariate integral, see Yunus and Khan (2009b) [15]. The density function of the bivariate noncentral chi-square distribution given above is a mixture of the bivariate central chi-square distribution of two central chi-square random variables with different degrees of freedom (see [19, 20]) with the probabilities from the Poisson distribution. Let $\rho^2 = \sum_{j=1}^{p} \frac{1}{p}\rho_j^2$, the mean correlation, where ρ_j is the correlation coefficient for any two different elements of the augmented vector $\left[\; n^{-\frac{1}{2}}M_{n_1}(0,\tilde{\beta}_2), n^{-\frac{1}{2}}M_{n_2}(\tilde{\beta}_1,0)\;\right]$ in equation (4.2).

We observe that when $\rho \neq 0$ and $\theta^{RT} \geq \theta^{UT}$, then (i) $\Pi^{PTT} < \Pi^{UT} \leq \Pi^{RT}$ for small (large) λ_1 and large (small) λ_2, (ii) $\Pi^{UT} \leq \Pi^{PTT} \leq \Pi^{RT}$ for small λ_1 and λ_2. These results are derived in Appendix B.

This confirms that the asymptotic size of the PTT is larger than that of the UT but less than that of the RT. For small and moderate values of λ_1 and λ_2, the asymptotic power of the PTT is larger than that of the UT but less than that of the the RT. But for large λ_1 or λ_2, the asymptotic power of the PTT may be smaller than that of the UT as well as the RT.

6 Illustrative example

For this illustrative example, we consider samples of size 100 from the multiple linear regression model in equation (1.1) with $p = 3$, $r = 1$ and $s = 2$. The random errors, e_i's $(i = 1, 2, \ldots, 100)$ are generated from the standard normal distribution using a code in R. Then, set $\beta_\nu = 1$ for $\nu = 1, 2, 3$. Let $c_{1i} = 1$ while c_{2i} and c_{3i} are 0 or 1 with 50% for each. In practice, often the normality assumption is not met due to the presence of contaminants in the collected data. In this example, to create contaminated observations, we randomly choose to replace $m(< n)$ of the n responses with some additive contamination, such that the contaminated responses X_i' is $X_i' = \beta_1 + \beta_2 c_{2i} + \beta_3 c_{3i} + d_i$ with d_i is generated from uniform distribution, $U[-5, -3.5]$ and $U[3.5, 5]$ with 50% for each. Only 10% contamination in the data is considered for simulation. For the contaminated data, the power functions of the UT, RT and PTT are calculated by equations (5.1), (5.2) and (5.4) using the Huber ψ-function, $\psi_H(U_i) = -k$ if $U_i < -k$, U_i if $|U_i| \leq k$, k if $U_i > k$, where $U_i = (X_i - \beta_1 - \beta_2 c_{2i} - \beta_3 c_{3i})/S_n$ with $S_n = MAD/0.6745$ and MAD is known as the mean absolute deviation. As suggested in many reference books (e.g [16, p.76]), the value of $k = 1.28$ is chosen because $k = 1.28$ is the 90th quantile of a standard normal distribution, so, there is a 0.8 probability that a randomly sampled observations will have a value between $-k$ and k. The estimate for σ_0^2 is taken to be $\sum \psi_H^2(U_i)/n$. For the estimation of γ, an R-estimate from the Wilcoxon sign rank statistic is used. The estimate of γ is the value of t such that $S(V_1, \ldots, V_n, t) = \sum_{i=1}^n \operatorname{sign}(V_i - t) a_n(R_{n_i}^+(t)) = 0$, where $R_{n_i}^+(t)$ is the rank of $V_i - t$ and $a_n(k) = k/(n+1), k = 1, \ldots, n$. Here, $V_i = \psi_H'(U_i)/S_n$ where $\psi_H'(U_i)$ is just the derivative of the Huber ψ-function.

Let $\boldsymbol{\lambda}_1 = [\ \lambda_1\ ]$ and $\boldsymbol{\lambda}_2 = [\ \lambda_2 \quad \lambda_3\ ]'$. Here, we set $\alpha_\nu = 0.05$ for $\nu = 1, 2, 3$ and consider all the cases when $\theta^{RT} \geq \theta^{UT}$. In Figure 1, the power of the UT, RT and PTT are plotted against λ_1 for selected values of $[\lambda_2, \lambda_3]$. As λ_1 grows large, power of all tests grow large too. Although the power of the UT and RT are increasing to 1 as λ_1 is increasing, the power of the PTT is increasing to a value that is less than 1. The analytical findings in the previous Section supports these graphical results.

Since the UT, RT and PTT are defined based on the knowledge of $\boldsymbol{\beta}_2 = [\ \beta_2 \quad \beta_3\ ]'$,

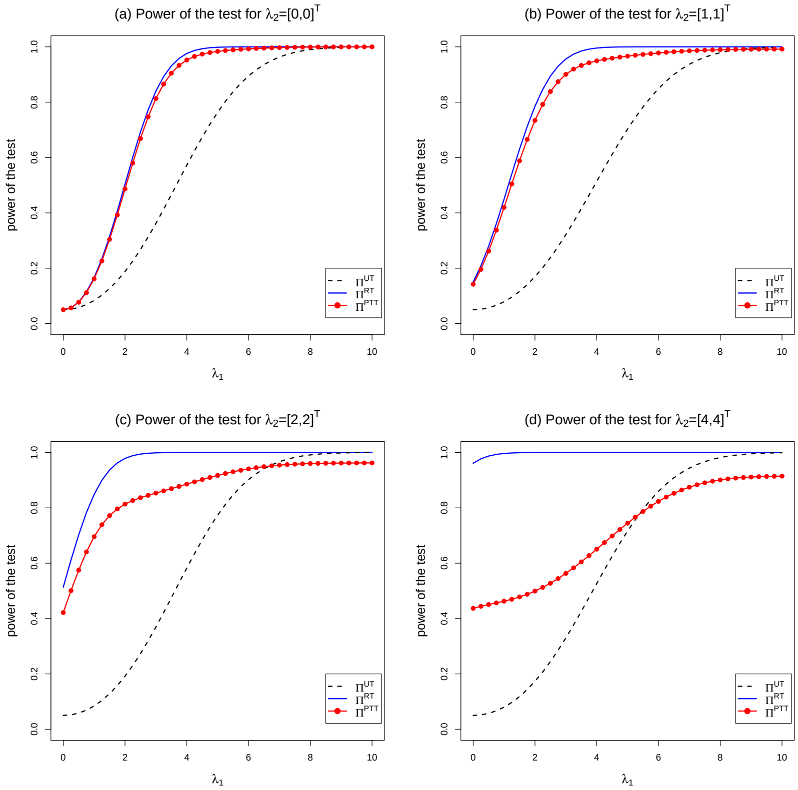

Figure 1: Graphs of power of the tests as a function of λ_1 for selected values of λ_2 and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha = 0.05$.

the size and power of each test are plotted against b such that $\lambda_2 = [\ b \quad b\ ]'$ in Figure 2. Figure 2 depicts that the RT has the largest power but also the largest size as b grows larger. On the contrary, the UT has constant smallest size regardless of the value of b but constant smallest power when $b < q$, where q is some positive value. From the observations, the PTT is a compromise in minimizing the size and maximizing the power when $b < q$. This is because it has smaller size than the RT but larger power than the UT. However, the PTT has the lowest power than the other tests when $b > q$. Although the prior information on the β_2 vector may be uncertain, there is a high possibility that the true values are not too far from the suspected values. Therefore, the study on the behaviour of the three tests when

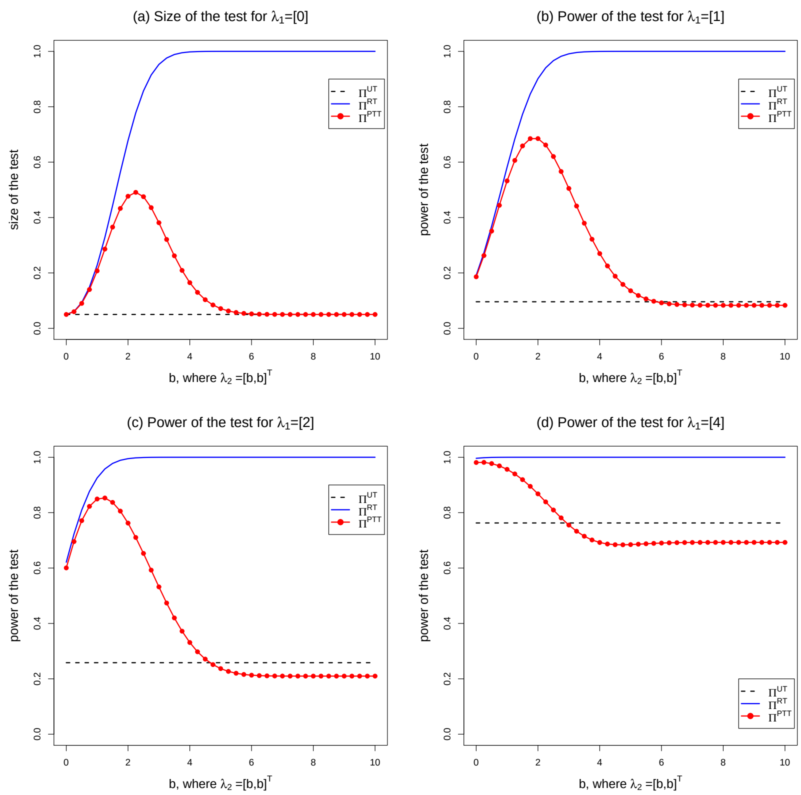

Figure 2: Graphs of power of the tests as a function of λ_2 for selected values of λ_1 and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha = 0.05$.

$b < q$ is more realistic. The properties of some of the proposed test statistics have already been studied and used to define preliminary test and shrinkage M-estimators by Ahmed et. al. (2006) [21].

7 Concluding Remarks

In this paper, the proposed M-tests do not depend on the assumptions on the distribution of the population. The asymptotic sampling distributions of the UT, RT and PT follow univariate noncentral chi-square distribution under the alternative hypothesis when the sample size is large. However, the sampling distribution of the PTT is a bivariate noncentral chi-square distribution as there is a correlation between the UT and PT. Note that there is no such correlation between the RT and PT. The new R code defined in Yunus and Khan (2009b) [15] is used for the computation of the distribution function of the bivariate noncentral chi-square distribution to evaluate the power function of the PTT .

The RT has the largest power among the three tests, but it also has the largest size. On the other hand the UT has smallest size, but it has the smallest power as well except when $\lambda_1 = n^{\frac{1}{2}}\beta_1$ or $\lambda_2 = n^{\frac{1}{2}}\beta_2$ is large. So, both UT and RT fail to achieve the highest power and lowest size simultaneously. The PTT has smaller size than the RT. It also has higher power than the UT, except for very large values of λ_1 or λ_2. Therefore if the prior information is not far away from the true value, that is, λ_2 is near 0 (small or moderate) the PTT has smaller size than the RT and higher power than the UT. Hence is it a better compromise between the two extremes. Since the prior information is coming from previous experience or expert knowledge, it is reasonable to expect λ_2 should not be too far away from 0, although it may not be 0, and hence the PTT demonstrate a reasonable domination over the other two tests in more realistic situation.

Acknowledgements

The authors thankfully acknowledge valuable suggestions of Professor A K Md E Saleh, Carleton University, Canada that helped improve the content and quality of the results in the paper.

A Appendix A

The following asymptotic results of [9, p.221]), [12, 14] are used in deriving the distribution of the proposed tests. For simplicity, we assume S is known or consider the nonstudentized M-estimator, so we omit condition M1 of [9, p.217] and let $S_n = S$ in equation (5.5.29) of [9, p.221]. Thus,

- Under $\beta_1 = a, \beta_2 = b$ where a and b are r and s dimensional column vectors of any real numbers, as n grows large,

$$\sup\{n^{-\frac{1}{2}}|M_{n_1}\{(a,b)+(t_1,t_2)\}-M_{n_1}(a,b)+n\gamma(Q_{11}t_1+Q_{12}t_2)|: \\ |t_1|\leq n^{-\frac{1}{2}}K,|t_2|\leq n^{-\frac{1}{2}}K\} \xrightarrow{p} 0, \tag{A.1}$$

$$\sup\{n^{-\frac{1}{2}}|M_{n_2}\{(a,b)+(t_1,t_2)\}-M_{n_2}(a,b)+n\gamma(Q_{21}t_1+Q_{22}t_2)|: \\ |t_1|\leq n^{-\frac{1}{2}}K,|t_2|\leq n^{-\frac{1}{2}}K\} \xrightarrow{p} 0. \tag{A.2}$$

- Under $\beta_1=0,\beta_2=0$, as n grows large,

$$n^{-\frac{1}{2}}\begin{pmatrix} M_{n_1}(0,0) \\ M_{n_2}(0,0)\end{pmatrix} \xrightarrow{d} N_p\left(\begin{bmatrix}0\\0\end{bmatrix},\sigma_0^2\begin{pmatrix}Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{pmatrix}\right), \tag{A.3}$$

where $N_p(\cdot,\cdot)$ represents a p-variate normal distribution with appropriate parameters and $K\in\Re$.

Proof of part (i) of Theorem 1. By equations (A.1) and (A.2), we find

$$n^{-\frac{1}{2}}M_{n_1}(0,\tilde{\beta}_2) = n^{-\frac{1}{2}}M_{n_1}(0,\beta_2)-n^{\frac{1}{2}}\gamma Q_{12}(\tilde{\beta}_2-\beta_2)+o_p(1) \text{ and} \tag{A.4}$$

$$n^{-\frac{1}{2}}M_{n_2}(0,\tilde{\beta}_2) = n^{-\frac{1}{2}}M_{n_2}(0,\beta_2)-n^{\frac{1}{2}}\gamma Q_{22}(\tilde{\beta}_2-\beta_2)+o_p(1) \tag{A.5}$$

under $H_0^\star$. Then, we obtain

$$n^{-\frac{1}{2}}M_{n_1}(0,\tilde{\beta}_2)=n^{-\frac{1}{2}}M_{n_1}(0,\beta_2)-n^{-\frac{1}{2}}Q_{12}Q_{22}^{-1}M_{n_2}(0,\beta_2)+o_p(1) \tag{A.6}$$

by equations (2.5), (A.4) and (A.5) after some simple algebra.

Further, the distribution of $n^{-\frac{1}{2}}M_{n_1}(0,\tilde{\beta}_2)$ under $H_0^\star$ is the same as the distribution of $n^{-\frac{1}{2}}M_{n_1}(0,0)-n^{-\frac{1}{2}}Q_{12}Q_{22}^{-1}M_{n_2}(0,0)$ under $H_0:\beta_1=0,\beta_2=0$ using equation (A.6) and the fact that the distribution of $M_{n_1}(a,b)$ under $\theta=a,\beta=b$ is the same as that of $M_{n_1}(\theta-a,\beta-b)$ when $\theta=0,\beta=0$, and similarly for $M_{n_2}(0,0)$ (c.f. [22, p.322]).

Therefore, utilizing equation (A.3), under $H_0^\star:\beta_1=0$ as $n\to\infty$, the proof of part (i) of Theorem 1 is completed. □

The proof for part (ii) of Theorem 1 is obtained in the same way as in part (i).

Proof of part (ii) of Theorem 2. Under $H_0:\beta_1=0,\ \beta_2=0$, with relation to (A.1) and (A.2),

$$\begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(\tilde{\beta}_1,0)\end{bmatrix} - \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(0,0)\end{bmatrix} + \begin{bmatrix} 0 \\ n^{\frac{1}{2}}\gamma Q_{21}\tilde{\beta}_1\end{bmatrix} \xrightarrow{p} \begin{bmatrix}0\\0\end{bmatrix}. \tag{A.7}$$

Note also that under H_0,

$$n^{-\frac{1}{2}}M_{n_1}(\tilde{\beta}_1,0) = n^{-\frac{1}{2}}M_{n_1}(0,0) - n^{\frac{1}{2}}\gamma Q_{11}\tilde{\beta}_1 + o_p(1) \tag{A.8}$$

and definition (2.6) reduce equation (A.8) to

$$n^{-\frac{1}{2}}Q_{21}Q_{11}^{-1}M_{n_1}(0,0) = n^{\frac{1}{2}}\gamma Q_{21}\tilde{\beta}_1 + o_p(1). \tag{A.9}$$

Therefore, under H_0, equation (A.7) becomes

$$\begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(\tilde{\beta}_1,0) \end{bmatrix} - \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(0,0) - n^{\frac{1}{2}}Q_{21}Q_{11}^{-1}M_{n_1}(0,0) \end{bmatrix} \tag{A.10}$$

$$= \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(\tilde{\beta}_1,0) \end{bmatrix} - \begin{bmatrix} I_r & 0 \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix} \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(0,0) \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{A.11}$$

Now utilizing the contiguity of probability measures (see [23, Ch.7]) under $\{K_n\}$ to those under H_0, the equation (A.11) implies that

$$\begin{bmatrix} n^{-\frac{1}{2}}M'_{n_1}(0,0) & n^{-\frac{1}{2}}M'_{n_2}(\tilde{\beta}_1,0) \end{bmatrix}'$$

under $\{K_n\}$ is asymptotically equivalent to the random vector

$$\begin{bmatrix} I_r & 0 \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix} \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(0,0) \end{bmatrix}$$

under H_0. But the asymptotic distribution of the above random vector under $\{K_n\}$ is the same as

$$\begin{bmatrix} I_r & 0 \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix} \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(-n^{-\frac{1}{2}}\lambda_1, -n^{-\frac{1}{2}}\lambda_2) \\ n^{-\frac{1}{2}}M_{n_2}(-n^{-\frac{1}{2}}\lambda_1, -n^{-\frac{1}{2}}\lambda_2) \end{bmatrix}$$

under H_0 by the fact that the distribution of $M_{n_1}(a,b)$ under $\theta = a, \beta = b$ is the same as that of $M_{n_1}(\theta - a, \beta - b)$ under $\theta = 0, \beta = 0$, and similarly for $M_{n_2}(0,0)$ (c.f. [22, p.322]).

Note that under H_0, with relation to (A.1) and (A.2),

$$\begin{aligned} \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(n^{-\frac{1}{2}}\lambda_1, n^{-\frac{1}{2}}\lambda_2) \\ n^{-\frac{1}{2}}M_{n_2}(n^{-\frac{1}{2}}\lambda_1, n^{-\frac{1}{2}}\lambda_2) \end{bmatrix} &= \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(0,0) \end{bmatrix} \\ &+ \begin{bmatrix} \gamma(Q_{11}\lambda_1 + Q_{12}\lambda_2) \\ \gamma(Q_{21}\lambda_1 + Q_{22}\lambda_2) \end{bmatrix} + \begin{bmatrix} o_p \\ o_p \end{bmatrix}. \end{aligned} \tag{A.12}$$

Hence, by equation (A.3), under H_0,

$$\begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(n^{-\frac{1}{2}}\lambda_1, n^{-\frac{1}{2}}\lambda_2) \\ n^{-\frac{1}{2}}M_{n_2}(n^{-\frac{1}{2}}\lambda_1, n^{-\frac{1}{2}}\lambda_2) \end{bmatrix} \to N_p\left(\begin{pmatrix} \gamma(Q_{11}\lambda_1 + Q_{12}\lambda_2) \\ \gamma(Q_{21}\lambda_1 + Q_{22}\lambda_2) \end{pmatrix}, \sigma_0^2\begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}\right). \tag{A.13}$$

Thus, as $n \to \infty$, the distribution of

$$\begin{bmatrix} n^{-\frac{1}{2}}M'_{n_1}(0,0) & n^{-\frac{1}{2}}M'_{n_2}(\tilde{\beta}_1, 0) \end{bmatrix}'$$

under $\{K_n\}$ is p-variate normal with mean vector

$$\begin{bmatrix} I_r & 0 \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix}\begin{pmatrix} \gamma(Q_{11}\lambda_1 + Q_{12}\lambda_2) \\ \gamma(Q_{21}\lambda_1 + Q_{22}\lambda_2) \end{pmatrix} = \begin{bmatrix} \gamma(Q_{11}\lambda_1 + Q_{12}\lambda_2) \\ \gamma(Q_{22} - Q_{21}Q_{11}^{-1}Q_{12})\lambda_2 \end{bmatrix}$$

and covariance matrix

$$\begin{bmatrix} I_r & 0 \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix}\sigma_0^2\begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}\begin{bmatrix} I_r & 0 \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix}' = \sigma_0^2\begin{pmatrix} Q_{11} & 0 \\ 0 & Q_2^\star \end{pmatrix}.$$

Since the two statistics $n^{-\frac{1}{2}}M_{n_1}(0,0)$ and $n^{-\frac{1}{2}}M_{n_2}(\tilde{\beta}_1, 0)$ are uncorrelated, asymptotically, they are independently distributed normal vectors. $\square$

Proof of part (i) of Theorem 2. Under $H_0 : \beta_1 = 0,\ \beta_2 = 0$, with relation to (A.1) and (A.2), (2.5), (A.9),

$$\begin{bmatrix} n^{-\frac{1}{2}}M'_{n_1}(0, \tilde{\beta}_2) \\ n^{-\frac{1}{2}}M'_{n_2}(\tilde{\beta}_1, 0) \end{bmatrix} - \begin{bmatrix} I_r & -Q_{12}Q_{22}^{-1} \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix}\begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(0,0) \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{A.14}$$

Now utilizing the contiguity of probability measures under $\{K_n\}$ to those under H_0, the equation (A.14) implies that

$$\begin{bmatrix} n^{-\frac{1}{2}}M'_{n_1}(0, \tilde{\beta}_2) & n^{-\frac{1}{2}}M'_{n_2}(\tilde{\beta}_1, 0) \end{bmatrix}'$$

under $\{K_n\}$ is asymptotically equivalent to the random vector

$$\begin{bmatrix} I_r & -Q_{12}Q_{22}^{-1} \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix}\begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(0,0) \\ n^{-\frac{1}{2}}M_{n_2}(0,0) \end{bmatrix}$$

under H_0. But the asymptotic distribution of the above random vector under $\{K_n\}$ is the same as

$$\begin{bmatrix} I_r & -Q_{12}Q_{22}^{-1} \\ -Q_{21}Q_{11}^{-1} & I_s \end{bmatrix} \begin{bmatrix} n^{-\frac{1}{2}}M_{n_1}(-n^{-\frac{1}{2}}\lambda_1, -n^{-\frac{1}{2}}\lambda_2) \\ n^{-\frac{1}{2}}M_{n_2}(-n^{-\frac{1}{2}}\lambda_1, -n^{-\frac{1}{2}}\lambda_2) \end{bmatrix}$$

under H_0. Then, equation (4.2) follows from equation (A.13) after some algebra. Since $n^{-\frac{1}{2}}M_{n_1}(0, \tilde{\beta}_2)$ and $n^{-\frac{1}{2}}M_{n_2}(\tilde{\beta}_1, 0)$ are uncorrelated, asymptotically, they are independently distributed normal vectors. □

B Appendix B

Write the second term on the right hand side of the equation (5.4) as

$$\int_{\chi^2_{r,\alpha_1}}^{\infty}\int_{\chi^2_{s,\alpha_3}}^{\infty} \phi^\star(w_1,w_2)dw_1dw_2 = [1 - H_r(\chi^2_{r,\alpha_1};\theta^{UT})][1 - H_s(\chi^2_{s,\alpha_3};\theta^{PT})], \tag{B.1}$$

where $H_r(\chi^2_{r,\alpha_1};\theta^{UT}) = \sum_{j=0}^{\infty}\sum_{\delta_1=0}^{\infty} \frac{(1-\rho^2)^{r/2}\Gamma(\frac{r}{2}+j)\rho^{2j}}{\Gamma(\frac{r}{2})j!}\gamma^\star\left(\frac{r}{2}+j+\delta_1, \frac{\chi^2_{r,\alpha_1}}{2(1-\rho^2)}\right)\frac{e^{-\theta^{UT}/2}(\theta^{UT}/2)^{\delta_1}}{\delta_1!}$

and $H_s(\chi^2_{s,\alpha_3};\theta^{PT}) = \sum_{k=0}^{\infty}\sum_{\delta_2=0}^{\infty} \frac{(1-\rho^2)^{s/2}\Gamma(\frac{s}{2}+k)\rho^{2k}}{\Gamma(\frac{s}{2})k!}\gamma^\star\left(\frac{s}{2}+k+\delta_2, \frac{\chi^2_{s,\alpha_3}}{2(1-\rho^2)}\right)\frac{e^{-\theta^{PT}/2}(\theta^{PT}/2)^{\delta_2}}{\delta_2!}$.

Note, $H_r(\chi^2_{r,\alpha_1};\theta^{UT}) \geq G_r(\chi^2_{r,\alpha_1};\theta^{UT})$ and $H_s(\chi^2_{s,\alpha_3};\theta^{PT}) \geq G_s(\chi^2_{s,\alpha_3};\theta^{PT})$. Equality is achieved when $\rho = 0$, or when $\lambda_1 = 0$ and $\lambda_2 = 0$.

Consider all the cases when $\theta^{RT} \geq \theta^{UT}$ and $\rho \neq 0$. So, using equation (B.1), we write equation (5.4) as

$$\Pi^{PTT}(\lambda_1,\lambda_2) \quad \leq \quad \Pi^{RT}(\lambda_1,\lambda_2)[1-\Pi^{PT}(\lambda_1,\lambda_2)] + \Pi^{UT}(\lambda_1,\lambda_2)\Pi^{PT}(\lambda_1,\lambda_2). \tag{B.2}$$

Equality in equation (B.2) is achieved when both λ_1 and λ_2 are 0. Obvious, $\Pi^{PTT}(\lambda_1,\lambda_2) < \Pi^{RT}(\lambda_1,\lambda_2) - \Pi^{PT}(\lambda_1,\lambda_2)v_2$ for $0 \leq v_2 < 1$, and it follows that $\Pi^{PTT}(\lambda_1,\lambda_2) \leq \Pi^{RT}(\lambda_1,\lambda_2)$ for any λ_1 and λ_2. Equality holds when both λ_1 and λ_2 are 0.

Rewrite equation (5.5) as

$$\begin{aligned} &\int_{\chi^2_{r,\alpha_1}}^{\infty}\int_{\chi^2_{s,\alpha_3}}^{\infty} \phi^\star(w_1,w_2)dw_1dw_2 \\ = \quad & 1 - H_r(\chi^2_{r,\alpha_1};\theta^{UT}) - H_s(\chi^2_{s,\alpha_3};\theta^{PT}) + \int_0^{\chi^2_{r,\alpha_1}}\int_0^{\chi^2_{s,\alpha_3}} \phi^\star(w_1,w_2)dw_1dw_2. \end{aligned} \tag{B.3}$$

When λ_2 is not large but not 0 and λ_1 is sufficiently large, the first term on the right hand side of the equation (5.4) becomes $G(\chi^2_{s,\alpha_3};\theta^{PT})$ because θ^{RT} is sufficiently large. The second and fourth terms on the right hand side of the equation (B.3) becomes 0 because θ^{UT} is large. Also, note that $H_s(\chi^2_{s,\alpha_3};\theta^{PT}) > G_s(\chi^2_{s,\alpha_3};\theta^{PT})$. So, $\Pi^{PTT} < \Pi^{UT} = 1$ for sufficiently large λ_1 and not so large λ_2 $(\neq 0)$.

Let $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$. When $\lambda_2 = 0$ and λ_1 is sufficiently large, the first term on the right hand side of the equation (5.4) becomes $1-\alpha$ because θ^{RT} is large. Both $H_r(\chi^2_{r,\alpha};\theta^{UT})$ and $\int_0^{\chi^2_{r,\alpha}} \int_0^{\chi^2_{s,\alpha}} \phi^\star(w_1,w_2)dw_1dw_2$ of the equation (B.3) become 0 while $H_s(\chi^2_{r,\alpha};\theta^{PT})$ becomes $1-\alpha$. So, $\Pi^{PTT} = \Pi^{UT} = 1$ when λ_1 is sufficiently large and $\lambda_2 = 0$.

In the same manner, we observe results given in (c)-(f) below

(a) When λ_2 $(\neq 0)$ is not large and λ_1 is sufficiently large, then, $\Pi^{PTT} < \Pi^{UT} = 1$.

(b) When $\lambda_2 = 0$ and λ_1 is sufficiently large, then, $\Pi^{PTT} = \Pi^{UT} = 1$.

(c) When λ_1 is not large but λ_2 is sufficiently large, then, $\Pi^{PTT} < \Pi^{UT} = 1$.

(d) When $\lambda_1 = 0$ and λ_2 is sufficiently large, then, $\Pi^{PTT} = \Pi^{UT} = \alpha$.

(e) When both λ_2 and λ_1 are sufficiently large, then, $\Pi^{PTT} < \Pi^{UT}$.

(f) When both λ_1 and λ_2 are 0, then $\Pi^{PTT} = \Pi^{UT} = \alpha$.

Also, we find from equations (5.4) and (B.1) that $\Pi^{PTT} > \Pi^{UT}$ if $\frac{\Pi^{RT}}{H_r(\chi^2_{r,\alpha};\theta^{UT})} > \frac{H_s(\chi^2_{s,\alpha};\theta^{PT})}{(1-\Pi^{PT})}$. We observe, when both λ_1 and λ_2 are 0, the $\Pi^{UT} = \Pi^{RT} = \Pi^{PTT} = \alpha$. As λ_1 grows larger and λ_2 fixed at 0, Π^{UT}, Π^{RT} and Π^{PTT} grow larger and approach 1 with $\Pi^{UT} \leq \Pi^{PTT} \leq \Pi^{RT}$.

When $\lambda_1 = 0$ and λ_2 is small $(\neq 0)$, we find $\Pi^{PTT}(0,\lambda_2) > \Pi^{UT}(0,\lambda_2) = \alpha$ if $\frac{\Pi^{RT}}{\alpha} > \frac{H_s(\chi^2_{s,\alpha};\theta^{PT})}{(1-\Pi^{PT})}$. As λ_1 grows larger and λ_2 fixed at some small positive vector, Π^{UT} and Π^{RT} grow large and approach 1, however, Π^{PTT} grows large and approaches a value less than 1. Note, $\Pi^{UT} \leq \Pi^{PTT} \leq \Pi^{RT}$ when $0 < \lambda_1 \leq q_1$, and $\Pi^{PTT} \leq \Pi^{UT} \leq \Pi^{RT}$ when $\lambda_1 > q_1$, where q_1 is some positive vector.

REFERENCES

[1] Khan, S. and Saleh, A.K.Md.E. (1997). Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate Student-t errors, *Biometrical Journal*, **39**, 131-147.

[2] Khan, S. and Saleh, A.K.Md.E. (2001). On the comparison of the pre-test and Stein-type estimators for the univariate normal mean, *Statistical Papers*, **42**, 451-473.

[3] Bechhofer, R.E. (1951). The effect of preliminary test of significance on the size and power of certain tests of univariate linear hypotheses. Ph.D. Thesis (unpublished), Columbia Univ.

[4] Bozivich, H., Bancroft, T.A. and Hartley, H. (1956). Power of analysis of variance test procedures for certain incompletely specified models, *Ann. Math. Statist.*, **27**, 1017 - 1043.

[5] Tamura, R. (1965). Nonparametric inferences with a preliminary test, *Bull. Math. Stat.*, **11**, 38-61.

[6] Saleh, A.K.Md.E. and Sen, P.K. (1982). Non-parametric tests for location after a preliminary test on regression, *Communications in Statistics: Theory and Methods*, **11**, 639-651.

[7] Saleh, A.K.Md.E. and Sen, P.K. (1983). Nonparametric tests of location after a preliminary test on regression in the multivariate case, *Communications in Statistics: Theory and Methods*, 12, 1855-1872.

[8] Lambert, A., Saleh A.K.Md.E. and Sen P.K. (1985). Test of homogeneity of intercepts after a preliminary test on parallelism of several regression lines: from parametric to asymptotic, *Comm. Statist.*, **14**, 2243-2259.

[9] Jurecčková, J. and Sen, P.K. (1996). *Robust Statistical Procedures Asymptotics and Interrelations.* John Wiley & Sons, US.

[10] Huber, P.J. (1981). *Robust Statistics.* Wiley. New York.

[11] Arnold, S.F. (2005). Nonparametric statistics. Summer School in Statistics for Astronomers and Physicists. Penn State University.

[12] Sen, P.K. (1982). On M-tests in linear models, *Biometrika*, **69**, 245-248.

[13] Yunus, R.M. and Khan, S. (2009a). Robust tests for multivariate regression model. Working Paper Series, SC-MC-0901, Faculty of Sciences, University of Southern Queensland, Australia

[14] Jurečková, J. (1977). Asymptotic relations of M-estimates and R-estimates in linear regression model, *Ann. Statist.*, **5**, 464-72.

[15] Yunus, R. M. and Khan, S. (2009b). The bivariate noncentral chi-square distribution - a compound distribution approach, Working Paper Series, SC-MC-0902, Faculty of Sciences, University of Southern Queensland, Australia.

[16] Wilcox, R.R. (2005). *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier Inc. US.

[17] Montgomery, D.C., Peck, E.A. and Vining G.G. (2001). *Introduction to linear regression analysis* (3rd ED.). Wiley. New York.

[18] Muirhead, R.J. (1982), *Aspects of multivariate statistical theory*, US: John Wiley & Sons.

[19] Gunst, R.F. and Webster, J.T. (1973). Density functions of the bivariate chi-square distribution, *Journal of Statistical Computation and Simulation A*, **2**, 275–288.

[20] Wright, K. and Kennedy, W.J. (2002). Self-validated computations for the probabilities of the central bivariate chi-square distribution and a bivariate F distribution, *J. Statist. Comput. Simul.*, **72**, 63–75.

[21] Ahmed, S.E., Hussein, A.A. and Sen, P.K. (2006). Risk comparison of some shrinkage M-estimators in linear models. *Nonparametric Statistics*, **18**, 401-415.

[22] Saleh, A.K.Md.E. (2006). *Theory of Preliminary test and Stein-type estimation with applications*. John Wiley & Sons. New Jersey.

[23] Hájek, J., Šidák, Z. and Sen, P.K. (1999). *Theory of Rank Tests.* Academia Press. New York.